Improving Situational Awareness in Emergencies through Crowd Supported Analysis of Social Media

Jakob Rogstadius,
Vassilis Kostakos
M-ITI, University of Madeira
9000-390 Funchal, Portugal
{jakob,vk}@m-iti.org

Jim Laredo, Maja Vukovic
IBM T.J. Watson Research Center
Hawthorne NY 10532, USA
{laredoj,maja}@us.ibm.com

Abstract. In this ongoing research project, we develop an information system that aims to improve situational awareness and shorten response times in emergency response situations. Through a combination of algorithmic and crowdsourcing techniques, the proposed system gathers, analyzes, organizes and then visualizes social media activity around an event in real-time and turns overwhelming streams of status updates into actionable pieces of information. This document is an extended abstract to the poster with the same name.

Social media in emergency response

Successful emergency response relies heavily on situational awareness, created from access to timely, accurate and relevant information about complex ongoing events. As a complement to traditional sources, researchers (Vieweg et al. 2010) and emergency response professionals (van der Vlugt and Hornery 2009) are now identifying social media as an emerging source of early breaking news, image and video footage, and an indicator of where to direct resources. However, existing information systems either fail to incorporate social media as a source, or do not meet the requirements imposed by use in crisis situations.

Algorithms vs. crowdsourcing

There are currently two main approaches for building real-time information systems. Purely automated news aggregators, such as EMM NewsBrief (Piskorki et al. 2008), already perform quite well at the task of gathering and clustering articles related to an event, including extracting metadata such as locations, people and quotes from the clusters. However, these systems offer generic approaches that are unable to gather and present knowledge in a manner tailored to the characteristics, needs and priorities of a specific event or disaster. Although social media aggregators exist, we are unaware of any that offer functionality and performance on a level similar to those for news.

Other systems more specialized for emergency use, such as Ushahidi (www.ushahidi.com), adopt an almost purely crowdsourced approach by relying on individuals to submit reports containing all necessary metadata; data which is then presented using default or in some cases event-adapted interfaces. While these systems are designed to be much more adaptive than the news aggregators, they are instead unable to integrate the vast but largely unstructured knowledge base related to a particular disaster that is social and traditional media.

Our contribution

The limitations of both fully automated and fully crowdsourced information processing systems motivate the need for solutions that combine the scalability of algorithmic computation, with the unique human capabilities to adapt to new situations, prioritize information, infer knowledge, estimate trust and question sources. Our proposed system (see poster) handles this by integrating crowdsourcing into an architecture of machine learning and NLP techniques, to analyze and structure social media content posted by microbloggers and service users during an event or disaster. The system is a work in progress and current functionality consists of topic tracking, message clustering, breaking news detection, an event timeline and drill-down functionality to read individual tweets.

References

Piskorski, J., Tanev, H., Atkinson, M. and van der Goot, E. (2008): 'Cluster-Centric Approach to News Event Extraction'. *In Proceeding of the 2008 conference on New Trends in Multimedia and Network Information Systems*, IOS Press, 2008, pp. 276-290.

Vieweg, S., Hughes, A., Starbird, K. and Palen, L. 'Microblogging during two natural hazards events: what Twitter may contribute to situational awareness'. In *Proc. CHI 2010*, ACM Press, 2010, pp. 1079-1088.

van der Vlugt, M., Hornery, A. (2009): 'Social Media helping Emergency Management – Final Report'. *NGIS Australia*, 2009. http://gov2.net.au/files/2009/12/Project-14-Final-Report.doc

Improving Situational Awareness in Emergencies through Crowd Supported Analysis of Social Media

Nobody here? Talk to me if you see me somewhere else.

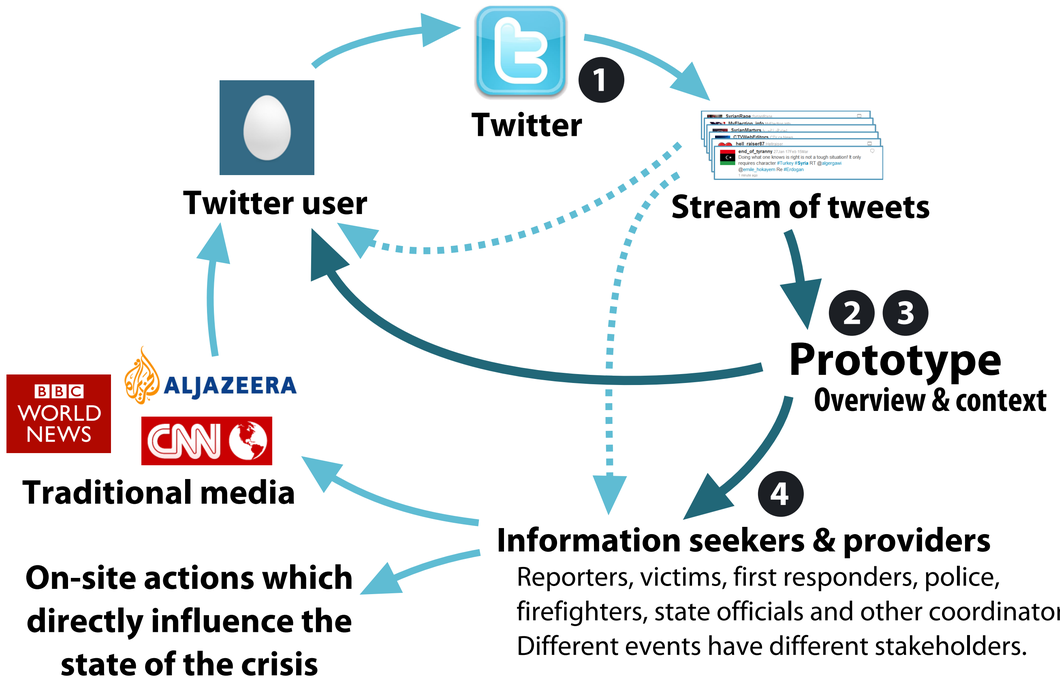
Jakob Rogstadius, Vassilis Kostakos
University of Madeira, {jakob,vk}@m-iti.org

Jim Laredo, Maja Vukovic
IBM T.J. Watson Research Center, {laredoj,maja}@us.ibm.com

ABSTRACT. We develop an information system that improves situational awareness and shortens response times in **emergency response** situations. The system gathers, analyzes, organizes and then visualizes **social media activity** around an event in **real-time** and turns overwhelming streams of status updates into actionable pieces of information.

CONTRIBUTION. The proposed system's novelty is that it integrates **crowdsourcing** into its architecture to analyze and structure social media content posted during the event or disaster by microbloggers and service users (including emergency response coordinators, victims and traditional media).

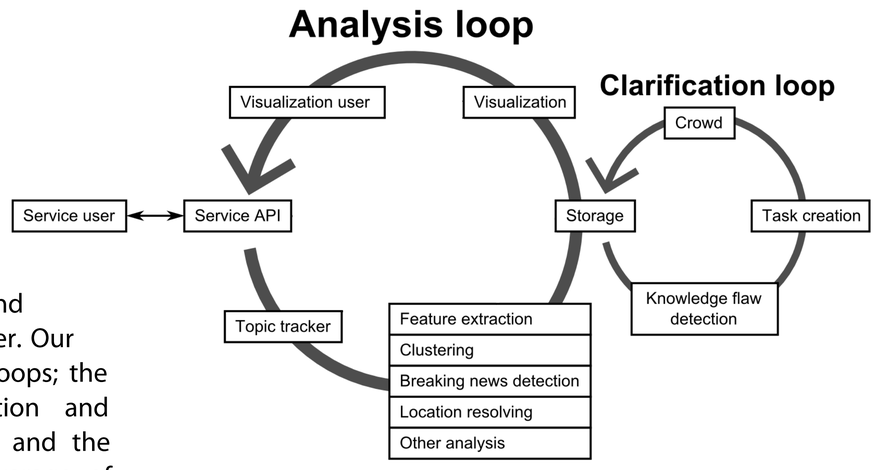
Flow of Information Surrounding an Emergency



Our Prototype

Architecture and Features

We work towards a system architecture (right) in which machine learning and NLP techniques work hand in hand with a crowdsourcing community to quickly and efficiently organize and analyze information around a disaster. Our design incorporates two feedback loops; the analysis loop structures information and enables better informed decisions; and the clarification loop improves the coverage of information regarding the event.

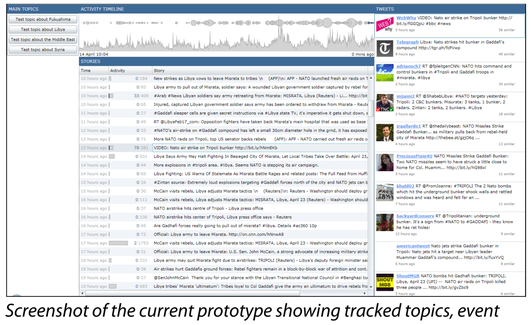


Implemented Features

- Topic tracking
- Breaking news
- Event timeline
- Drill-down to individual tweets
- Source ranking

Work in Progress

- Crowd integration
- Location and entity tagging for stories
- External in-depth articles
- Images & video
- Trust inference



Screenshot of the current prototype showing tracked topics, event timeline, detected stories and tweets in a selected story.

1 Social media, e.g. Twitter, has in numerous crises proven to contain timely and detailed information about ongoing events. However, it is difficult to gain an overview of the information, in particular for decision makers who may already be under great cognitive load.

2 We employ data mining techniques to cluster similar messages into stories. This helps preventing information overload when hundreds of messages are posted per minute.

3 Placing stories on a timeline (see screenshot) lets users go back in time to see what lead up to current events. In addition, an activity graph makes it easy to find eventful time periods and to compare current and past activity.

4 By giving decision makers improved situational awareness, they are empowered to make better informed and more timely decisions and to take better actions which directly influence the state of the crisis.